

09/894,898

**REMARKS**

In view of the following discussion, the Applicant submits that none of the claims now pending in the application is anticipated under the provisions of 35 U.S.C. § 102 or made obvious under the provisions of 35 U.S.C. § 103. Thus, the Applicant believes that all of these claims are now in allowable form.

**I. REJECTION OF CLAIMS 1, 2, 8-26, 28, 30 AND 32-35 UNDER 35 U.S.C. § 102**

The Examiner has rejected claims 1, 2, 8-26, 28, 30 and 32-35 in the Office Action as being anticipated by the Brown et al. patent (US patent 5,719,997, issued on February 17, 1998, hereinafter "Brown"). In response, the Applicant has amended independent claims 1, 11, 18, 34 and 35, from which claims 2, 8-10, 12-17, 19-26, 28, 30 and 32-33 depend, in order to more clearly recite aspects of the present invention.

Brown teaches a speech recognition system that uses an evolutionary grammar to recognize an input speech signal in real time. In particular, Brown teaches that as speech recognition processing begins, only a portion of a system grammar (*i.e.*, a vocabulary comprising a plurality of interrelated words) is implemented for recognition purposes. As more of the speech signal is received by the system and as processing proceeds, additional portions of the grammar network (*i.e.*, additional words or vocabulary) are implemented as necessary. In other words, a single system grammar is assembled, piece-by-piece, as the speech signal is received.

The Examiner's attention is directed to the fact that Brown fails to disclose or suggest the novel invention of acquiring at least part of the information necessary to generate a grammar and one or more related subgrammars (including, for example, a word subgrammar, a phone subgrammar and a state subgrammar) from a remote computer, as claimed in Applicant's independent claims 1, 11, 18, 34 and 35. Specifically, Applicant's claims 1, 11, 18, 34 and 35 positively recite:

1. A method for allocating memory in a speech recognition system comprising the steps of:

acquiring a first set of data structures that contain a grammar, a word subgrammar, a phone subgrammar and a state subgrammar, each of the subgrammars related to the grammar, wherein the first set of data structures is generated by the

09/894,898

speech recognition system using information provided at least in part by a remote computer;

acquiring a speech signal;  
performing a probabilistic search using the speech signal as an input, and using the grammar and the subgrammars as possible inputs; and  
allocating memory for one of the subgrammars when a transition to that subgrammar is made during the probabilistic search. (Emphasis added)

11. In a speech recognition system, a method for recognizing speech comprising the steps of:

acquiring a first set of data structures that contain a grammar, a word subgrammar, a phone subgrammar and a state subgrammar, each of the subgrammars related to the grammar, wherein the first set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer;

acquiring a speech signal;  
performing a probabilistic search using the speech signal as an input, and using the grammar and the subgrammars as possible inputs;  
allocating memory for one of the subgrammars when a transition to that subgrammar is made during the probabilistic search; and  
computing a probability of a match between the speech signal and an element of the subgrammar for which memory has been allocated. (Emphasis added)

18. In a speech recognition system, a method for recognizing speech comprising the steps of:

acquiring a first set of data structures that contain a top level grammar and a plurality subgrammars, each of the subgrammars hierarchically related to the grammar and to each other, wherein the first set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer;

acquiring a speech signal;  
performing a probabilistic search using the speech signal as an input, and using the top level grammar and the subgrammars as possible inputs;  
allocating memory for specific subgrammars when transitions to those specific subgrammars are made during the probabilistic search; and  
computing probabilities of matches between the speech signal and elements of the subgrammars for which memory has been allocated. (Emphasis added)

34. A method for allocating memory in a speech recognition system comprising the steps of:

acquiring a set of data structures that contain a grammar and one or more subgrammars related to the grammar, wherein the set of data structures is generated by

09/894,898

the speech recognition system using information provided at least in part by a remote computer;

acquiring a speech signal;  
performing a probabilistic search using the speech signal as an input, and using the grammar and the subgrammars as possible inputs; and  
allocating memory for a selected one or more of the subgrammars when a transition to the selected subgrammar is made during the probabilistic search.  
(Emphasis added)

35. In a speech recognition system, a method for recognizing speech comprising the steps of:

(a) acquiring a set of data structures that contain a grammar and one or more subgrammars related to the grammar, wherein the first set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer;

(b) receiving spoken input;  
(c) using one or more of the data structures to recognize the spoken input;  
(d) while the speech recognition system is operating, acquiring a second set of data structures that contain a second grammar and one or more subgrammars related to the second grammar; and  
(e) repeating steps (b) and (c), using the second set of data structures in step (c). (Emphasis added)

Applicant's invention is directed to a method for allocating memory in a speech recognition system. Conventional speech recognition systems require a great deal of memory in order to accommodate and process large vocabularies. These systems typically compile, expand, flatten and optimize all grammars contained in a system vocabulary into a large, single-level data structure that must be stored in memory before the speech recognition system can operate. Such techniques substantially restrict the capabilities of speech recognition systems that operate on limited memory and processing power, such as portable speech recognition systems.

The present invention provides a method for speech recognition in which memory is allocated to a particular system subgrammar when a transition is made to that subgrammar during a probabilistic search. A system vocabulary has a hierarchical data structure including at least one top-level grammar (e.g., "Days of the Week") and at least one subgrammar within that top-level grammar such as a word subgrammar (e.g.,

09/894,898

Monday, Tuesday, Wednesday, etc.), a phone subgrammar (e.g., /m/, /ah/, /n/, /d/, /ey/, etc.) and a state subgrammar (e.g., comprising Hidden Markov Models). When the system receives a speech signal for processing, the speech signal is input, along with the (unexpanded) top-level grammar and one or more subgrammars, into a probabilistic search. When a transition is made to a particular subgrammar during the probabilistic search, memory is allocated to the subgrammar, which may then be expanded and evaluated to assess the probability of a match between the speech signal and an element in the subgrammar. In this manner, memory is conserved and allocated only to portions of the system vocabulary that are currently needed for speech processing. In addition, at least part of the information used to generate the top-level grammar and/or the related subgrammars may be provided by a remote computer or server, to further conserve the memory required to operate the speech recognition system (which may be implemented, for example, in a portable device).

In contrast, Brown teaches a method in which a speech recognition system possesses an entire grammar, but merely instantiates selected portions of the grammar over time, e.g., as more of the incoming speech signal is received. Thus, Brown fails to anticipate or make obvious Applicant's invention.

Specifically, as the Examiner concedes on page 14 of the Office Action, "Brown does not teach that the set of data structures [i.e., the top-level grammar and associated subgrammars] is sent by a communication channel by a remote computer, or selected thereby or that the set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer" (emphasis added). In other words, Brown does not teach, show or suggest the desirability of distributing the top-level grammar and related subgrammars in order to further conserve memory at the speech recognition system. Therefore, the Applicant submits that independent claims 1, 11, 18, 34 and 35 fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

Dependent claims 2, 8-10, 12-17, 19-26, 28, 30 and 32-33 depend from claims 1, 11 and 18 and recite additional features therefore. As such, and for at least the same reasons set forth above, the Applicant submits that claims 2, 8-10, 12-17, 19-26, 28, 30

09/894,898

and 32-33 are not anticipated by the teachings of Brown. Therefore, the Applicant submits that dependent claims 2, 8-10, 12-17, 19-26, 28, 30 and 32-33 also fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

## **II. REJECTION OF CLAIMS 3-7, 16-17 and 36 UNDER 35 U.S.C. § 103**

The Examiner rejected claims 3-7, 16-17 and 36 under 35 U.S.C. §103(a) as being unpatentable over Brown in view the Ehsani et al. application (U.S. Publication No. 2002/0032564, published March 14, 2002, hereinafter "Ehsani"). In response, the Applicant has amended independent claims 1 and 11, from which claims 3-7 and 16-17 respectively depend, as discussed above in order to more clearly recite aspects of the invention. The rejection with respect to claim 36 is respectfully traversed.

Brown has been discussed above. Ehsani teaches a method for creating grammar networks for use in natural language voice user interfaces (NLVUIs). Valid phrases are extracted from a text corpus and clustered into classes to create a "thesaurus" of fixed word combinations that represent different ways of saying the same thing. In this way, anticipated user responses can be expanded into alternative linguistic variants.

The Examiner's attention is directed to the fact that Brown and Ehsani (either singly or in any permissible combination) fail to disclose or suggest the novel invention of acquiring at least part of the information necessary to generate a grammar and one or more related subgrammars (including, for example, a word subgrammar, a phone subgrammar and a state subgrammar) from a remote computer, as claimed in Applicant's independent claims 1, 11 and 36. Independent claims 1 and 11 have been recited above. Applicant's independent claim 36 positively recites:

36. In a speech recognition system, a method for recognizing speech comprising the steps of:

- (a) acquiring from a first remote computer a set of data structures that contain a grammar and one or more subgrammars related to the grammar;
- (b) receiving spoken input;
- (c) using one or more of the data structures to recognize the spoken input;
- (d) while the speech recognition system is operating, acquiring a second set of data structures from the first remote computer or from a second remote computer, the

09/894,898

second set of data structures containing a second grammar and one or more subgrammars related to the second grammar; and

(e) repeating steps (b) and (c), using the second set of data structures in step (c). (Emphasis added)

As recited in the preceding claim, Applicant's invention teaches a method for speech recognition in which memory is allocated to a particular system subgrammar (e.g., a word, phone or state subgrammar) when a transition is made to that subgrammar during a probabilistic search. Memory allocation allows the subgrammar to be expanded and evaluated to assess the probability of a match between an input speech signal and an element in the subgrammar. In this manner, memory is conserved and allocated only to portions of the system vocabulary that are currently needed for speech processing. In addition, at least part of the information used to generate the top-level grammar and/or the related subgrammars may be provided by a remote computer or server, to further conserve the memory required to operate the speech recognition system (which may be implemented, for example, in a portable device).

In contrast, neither Brown nor Ehsani teaches or suggests this novel approach. Neither Brown nor Ehsani teaches that the set of data structures (i.e., the top-level grammar and associated subgrammars) is sent by a communication channel by a remote computer, or selected thereby or that the set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer, as positively claimed by the Applicant in claims 1, 11 and 36.

The portions of Ehsani that the Examiner cites to support this limitation do not, in fact, teach that a speech recognition system acquires a grammar, or information necessary to generate a grammar, from a remote computer. Specifically, the cited portions of Ehsani teach that a voice telephony server may be accessed remotely by a user using a telephone or other handheld device. The voice telephony server guides the user through a series of voice-driven interactions that lets the user complete an automated transaction (e.g., getting information, accessing a database or making a purchase). In other words, Ehsani teaches that the speech signal or input to be

09/894,898

processed may be provided remotely (e.g., via a telephone or handheld device). However, the grammar used to process the remotely provided speech signal resides locally, on the voice telephony server. Therefore, the grammar used to process the speech signal is not provided by a remote computer or server. Therefore, the Applicant submits that independent claims 1, 11 and 36 fully satisfy the requirements of 35 U.S.C. §103 and are patentable thereunder.

Dependent claims 3-7 and 16-17 depend, either directly or indirectly, from claims 1 and 11 and recite additional features thereof. As such and for at least the same reasons set forth above, the Applicant submits that claims 3-7 and 16-17 are also not made obvious by the teachings of Brown in view of Ehsani. Therefore, the Applicant submits that dependent claims 3-7 and 16-17 also fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

### III. CONCLUSION

Thus, the Applicant submits that all of the presented claims now fully satisfy the requirements of 35 U.S.C. §102 and §103. Consequently, the Applicant believes that all of these claims are presently in condition for allowance. Accordingly, both reconsideration of this application and its swift passage to issue are earnestly solicited.

If, however, the Examiner believes that there are any unresolved issues requiring the issuance of a final action in any of the claims now pending in the application, it is requested that the Examiner telephone Mr. Kin-Wah Tong, Esq. at (732) 530-9404 so that appropriate arrangements can be made for resolving such issues as expeditiously as possible.

8/29/05  
Date

Moser, Patterson & Sheridan, LLP  
595 Shrewsbury Avenue  
Shrewsbury, New Jersey 07702

Respectfully submitted,

Diana J. Rea  
Diana J. Rea, Attorney  
Reg. No. 54,938  
(732) 530-9404